Annexure-5

SENTIMENT ENRINCHED CROSS DOMAIN RECOMMENDER SYSTEM

SHORT SYNOPSIS

Submitted in partial fulfillment of the requirement of the degree of

DOCTOR OF PHILOSOPHY

to

Manav Rachna University

by

Nikita Taneja

Registration number: 2K17-PFE-007

Under the supervision of

Dr. Hardeo K Thakur
Associate Professor
Department of Computer Science & Technology
Faculty of Engineering

MANAV RACHNA vidyanatariksha
MANAV RACHNA UNIVERSITY
(FORMERLY MANAV RACHNA COLLEGE OF ENGINEERING
NAAC ACCREDITED 'A' GRADE INSTITUTION)
Declared as State Private University under section 2f of the UGC act, 1956

MANAV RACHNA
vidyanatariksha

MANAV RACHNA UNIVERSITY
(FORMERLY MANAV RACHNA COLLEGE OF ENGINEERING
NAAC ACCREDITED 'A' GRADE INSTITUTION)
Declared as State Private University under section 2f of the UGC act, 1956

Date: 28th Oct 2021

DECLARATION BY THE SCHOLAR

I hereby declare that the short synopsis in the area of Sentiment Enriched Cross Domain Recommender System being submitted in partial fulfillment of the requirement for the degree of Doctor of Philosophy in department of **Computer Science and Technology** under **Faculty of Engineering** of Manav Rachna University, Faridabad during the academic year **2021.**

It is the work that is expected to be carried out under the supervision of **Dr. Hardeo Kumar Thakur.** Further, it is certified that short synopsis is within the plagiarism limit and copy of the same is attached as per University norms.

Nikita. 28/10/21

(Signature of the Scholar)

Ms. Nikita Taneja
MRUCSDN012K160395

Hardeo

Signature of the Supervisor

Dr. Hardeo Kumar Thakur
Associate Professor
Department of Computer Science and Technology
Manav Rachna University

1. Introduction

Recommendation systems have a significant role to play in how we surf the internet e.g. recommendation given by YouTube decides which video we will see next. This could also have some negative consequences e.g. if we recommend users only one type of content, they can be easily manipulated [1].

Recommender system is:

(a) Data analytics software and main area in information filtering field and data mining.

(b) Subset of information filtering techniques filters relevant information from irrelevant data.

(c) Software useful in content, product or services (such as movies, books, music, digital products, shopping, cloths, web sites, friends group etc.) abstraction.

Recommender system helps in decision making process and handling with information overload problem. In the websites related to e-commerce such as Amazon and many more, to exploit the evaluations of the user would be highly significant regarding many kinds of products so as to produce a model which is general for the preferences of user. As a matter of fact, there could be many correlations as well as dependencies which can exist between different domain preferences and despite of each item treatment independently such as music, electrical products etc. the knowledge of the user is obtained inside a domain which is able to be shifted and several other domains can be used to exploit it. [2].

The main cross domain recommendation goal furthermore is not this but however a join offer could be given by a system, in multiple domains the personalized recommendations of items for example a movie recommendation, CDs music and books, any video game which is related somehow to the movie. In the same way in the applications of tourism, a cultural event can also be suggested to a client who is interested to get a room booking in a hotel recommended, also in the e learning mechanism with bibliographic references to the student which to video literature are related and has been recommended recently. Joint recommendations are also offered by some systems in different domains for an item, but collectively speaking, in one domain so as to build a recommendation, the user preferences are only exploited by them on the domain which is targeted.

One of the research topics which is seen to be emerging now days is the cross domain recommendation. On the cross domain recommendation, Winoto and Tang presented one of

the initial studies. In the work which they presented, the three important issues were identified by the authors and these issues are supposed to be investigated.

(a) Global correlations existence and their verification for the items in various domains and user preferences.

(b) Model designing which are able to make use of the user based preferences on the domain source for the user preference prediction on the domain targeted.

(c) For cross domain recommendations the development of the appropriate evaluations.

Irrespective of the work and advantages one gets, there is an innovation in this research topic and most of the part of this topic is still left unexplored. To the best of our information hence, for a cross domain problem an agreed definition in literature is supposed to offer a joint, not far emerging system with no existing approaches and their general classification. How to exploit information and this issue from lot of other domains to offer suggestion from the distinct

perspectives has been addressed in many areas which include the retrieval of information, management of knowledge, user modeling, and learning via machine.

Progress or achievement of any recommender system is dependent generally upon two factors:

(a) Recommender algorithmic accuracy

(b) User characteristics and number of products in a given system.

Every recommender system "should follow different recommendation process depends upon the types of information they use to make recommendations. Two approaches are used in recommender systems that are Personalized Recommendation Systems and Non-personalized recommendation systems. Non-Personalized recommender systems also called a system of user based recommender which is able to produce the suggestions depending on what the other customers have said about the product, recommendations are not make by using the personal user preferences.

1.1 RECOMMENDATION SYSTEM TECHNIQUES

There are several categories of the recommender system for its classification and is generally based on the data which is used to suggest the product. On the other hand, in the literature following three categories are followed generally:

(i) RS which is based on content

(ii) Collaborative-Based RS

(iii) Hybrid RS

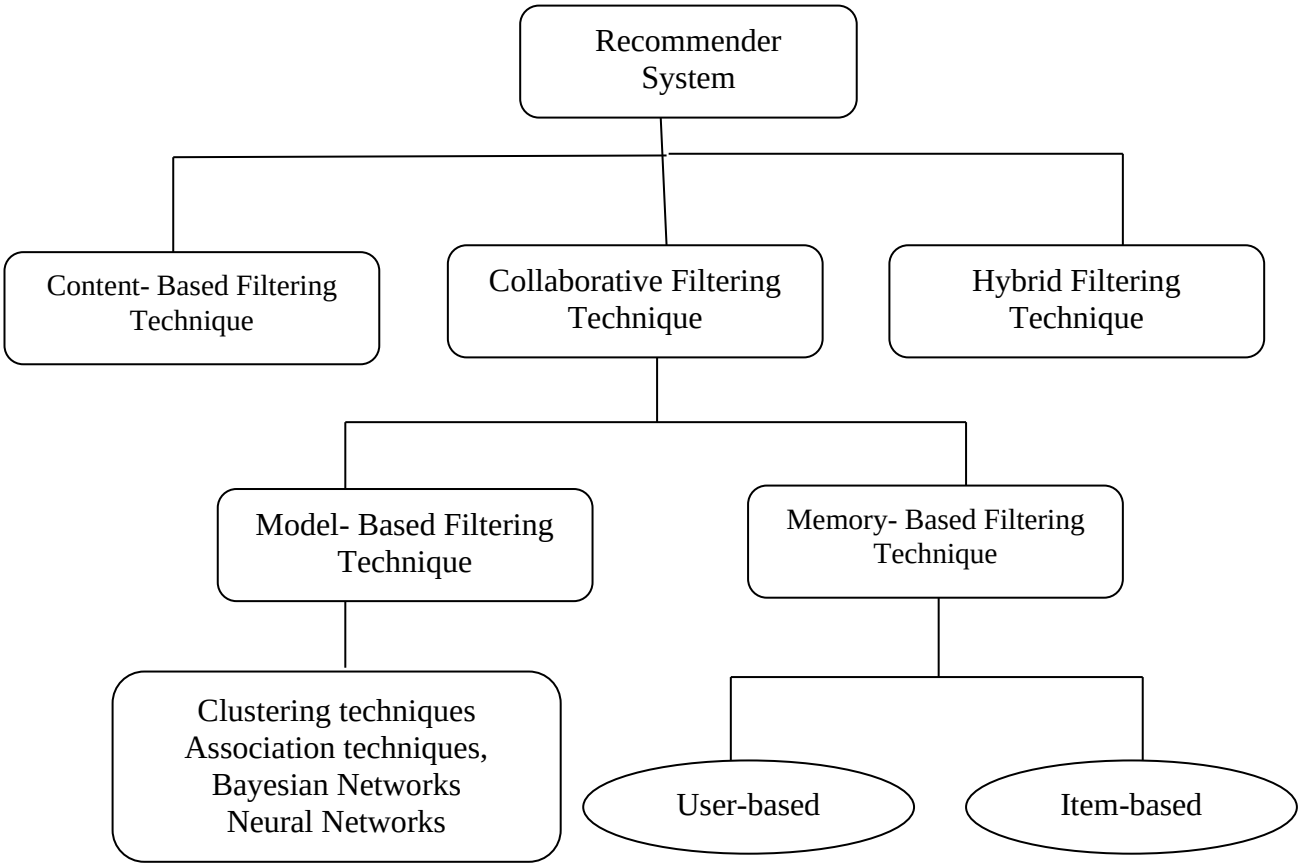


Figure 1: Classification of recommendation techniques

1.2 Cross Domain Recommendation System, main issues and challenges:

1.2.1 Synonymy

When an item or product is represented with more than two names there is synonymy which arises, this also is done by the entries which have the same meanings. The cases like this, whether the terms represent different items or same item is not able to be identified by the recommender. The CF [3] approach which is based on memory for example will be able to treat differently the comedy movie or a comedy film.

1.2.2 Privacy

To the systems of recommender, feeding the personal information often ends up in services with better recommendation, but this at the same time can result in the concerns regarding data security and privacy. Into the system recommender the users are seen to be reluctant to

feed data which are more prone to the concerns regarding data safety and privacy. Hence, the system of recommender which is either CB or CF must be able to make a relationship of strong trust with the customers; on the other hand the recommenders of CF are seen to be highly prone to the issues such as data security and privacy [4].

1.2.3 Limited Content Analysis and Overspecialization

The recommendations which are based on content are seen to be relying on the general content regarding products and customers supposed to be processed by the techniques for the retrieval of data. The content when available in a limited form can cause the over specialization kind of problems. By their subjective attributes in this case the products are represented and in this case the selection of item is generally dependent on their biased characteristics.

1.2.4 Sparsity

The data of larger size which is available regarding the catalogue of products and the users disinclination for the item rating often makes up a matrix of dispersed profile and this can lead to recommendations which are lesser accurate. In the systems of CF the sparse rating often makes it tough to make prediction which is accurate regarding the products. To recommend the items the users of CF are nearest to neighbors and ratings are often lesser which make it hard to compute and estimate the surrounding parties. In context aware recommenders the problem can be intense as multidimensional vectors are used by them and to provide accurate recommendations is also tough to be provided for the items which are very few rated. In order to meet the needs of these circumstances, one can make use of various approaches which include the models of multidimensional recommendations, the techniques of SVD [5,6], the filtering on the basis of demographics and CF algorithmic boosting of the content.

1.2.5 Scalability

In the neighbor nearest algorithms the rate of growth demonstrates the association which is linear with the product number and user number. For a typical recommender it becomes difficult to process a data which is of large scale. Amazon.com is an example of it which suggests products over 18 million and their customers are over 20 million. [8].

1.2.6 Latency Problem

The latency problem is faced by the recommenders of CF when there is frequent addition of new products to the database in a case where the suggestion is made by the recommenders on the items already rated as items which are added newly and are not yet given any rating. Making use of filtering with CB the waiting time can be reduced but at the same time there is a chance of introduction of over specialization.

1.2.7 Evaluation and the Availability of Online Datasets

The recommender system evaluation is also supposed to determine the other aspects such as quality. The evaluation criteria and its design along with the suitable evaluation selection metrics are regarded as problems in the systems of recommenders. The systems of recommenders which are more traditional help in the estimation of the system outcomes and algorithmic evaluations and this is done by considering the test sets which are obtained by the division of datasets and some metrics are applied such as precision, MAE, F measure for the proper calculations.

1.2.8 Context-Awareness

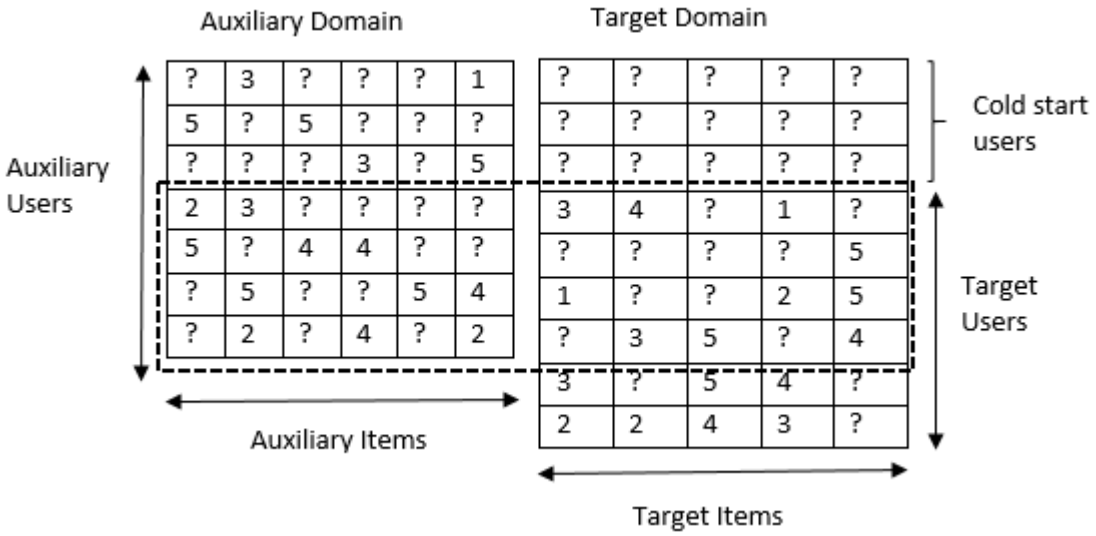
Operation point of view suggests that there is an aggregation of the context awareness and different categories which show the adjustments where there is deployment of the recommender, for example, the location used at present or current location, the recent activity and the period of time. The upcoming systems of recommenders as envisioned are supposed to make use of data which is contextual and is collected via infrastructure of cellular services and user's history both long and short term are included in it along with the current location, calendar entries and data which a user saves for the social media networking sites.

1.2.9 Cold Start Problem

When the new users enter the system this problem takes place and also in the case when into the catalogue more items are added. In any of the cases as mentioned, one cannot predict the taste of the new users also no one can suggest the ratings of new products which users purchase resulting in recommendations which are lesser accurate. In many ways the problem of cold start can be solved which include, a) user is requested to give rating to the product, b) explicitly asking a user to mention aggregate taste, c) recommending the products which are

based on user on the information or data which is demographically collected. To know about the location and the zip code the user demographic information can be used along with the new user interactions with the recommended system so as to suggest the products based on their user ratings. The ratings are offered by the users of the same product having data of similar demographics. There are however some domains whereby the items which are sleepers and are considered good but due to some reasons they are not given any rating. Using metadata the sleepers can be well handled, this can also be done by content based methodology making use of popularity of an item, entropy of an item and personalization of the user.

In actual, the users of cold start for the domain of a product might contain ratings in the domain of other product. For example , in one of the largest Chinese B2C e-commerce website Jingdong5, most users tend to buy electronic products while are much less interested in purchasing other types of products such as books and foods. On the contrary, in the e-commerce website, users prefer to purchase books and will probably not purchase its electronic products. Thus, we focus on studying the recommendation of the cross domain for the users of cold start and setting problem which in this paper is studied and is mentioned below in the diagram [9].



We differentiate item domains based on the item type level addressed by the work. For example, movies and books represent different domains, but horror movies and comedy movies belong to the same domain. One can see from Fig. above that in our problem setting the users of cold start in the auxiliary domains have their ratings which from the previous work is different and the auxiliary domain data is assumed which is denser relatively as

compared to the data of the targeted domain in the case where users of cold start are not considered. In both domains the users who have their ratings are known as linked customers or linked users and dashed red boxes are marked in their data rating as mentioned above in the figure. . Linked users serve as a bridge for our model to transfer knowledge across domains. To make cold start users in the targeted domains is always challenging in order to make any recommendation. In various item domains at first the ratings are normally sparse and hence the way to model the unique features better for the different domains users become highly significant. In the second step, for users of cold start no data rating is done in the targeted domains and there is a correlation in the behavior of two users in different domains. Due to this, the type of information or knowledge which is supposed to be spread and the ways of transferring the knowledge across certain domains is supposed to be an open concern.

1.3 Problem Formulation

Collaborative Filtering (CF) is a superior algorithmic recommendation which at present are used widely. In the recommender system the CF is especially designed so that it can predict the ratings which are missed for a customer on the product based from the likeminded users of the collected ratings or the like products. There is much simplicity in CF and no information configuration is needed for the users and there are no particular requirements and object recommendations. There is much more effectiveness in CF and multiple recommendations are supposed to be made by it. There are such beneficial aspects of CF but regardless of it the performance recommendations are supposed to be seriously degraded when there is sparseness in the data which is observed [10]. In order to deal with the problems of sparseness the CF methods which are much improved are suggested which on the single domain are based. For the targeted domains these methods are subject to the quality of data. In case where the targeted data is extremely sparse they may not be valid always. Indeed, the big data era at present, in the associated domains we have many data for the targeted domains. Why it is not recommended to make use of this data? A cross domain recommendation therefore is investigated by us combining the data which is relevant from various other domains with the genuine targeted data for the recommendations and their improvement. The process of Transfer learning for the CF recommendation can be used especially. Useful information or knowledge during the transfer learning will be obtained via

data which is auxiliary and is shifted later to the targeted data so that the problems or concerns of its thinness or sparseness can effectively be addressed. .[11][12][13][14][15] on the other hand, for the cross domain CF recommendations the traditional learning models are supposed to face some key concerns which are mentioned below:

- For the transferring of the homogenous user feedback they are often restricted. On the other hand, the feedback of heterogeneous users in reality is common particularly in the era where data is bigger.
- Different domains are normally supposed by them and common pattern of latent rating is shared by them which on the co-clustering of user item is based on. Indeed, the domains which are linked sometimes also lack the sharing of patterns with common latent rating, and inside the associated domains the common latent ratings which can also cause degradation of the performance recommendation.

- There is sparseness in the targeted data, more common useful knowledge as expected can be shifted from the auxiliary to targeted data. Making use of the latent factors exclusively which from the auxiliary data are obtained can end up in the generation of positive information in the case where there is insufficiency in the targeted data.

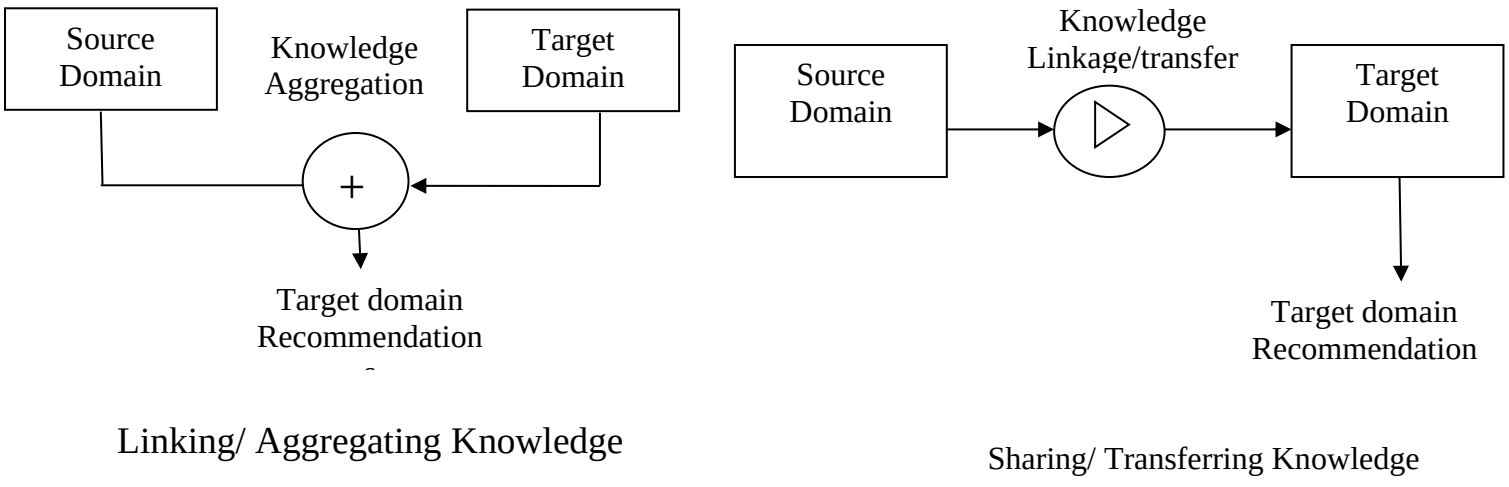


Figure 3: Transfer Learning in Collaborative Filtering

1.4 Motivation

The problems of sparseness are faced by the cross domain collaborative filtering and this is done by the shifting of rating information in various existing domains. On the other hand, the

from auxiliary domains the transfer knowledge is often nontrivial. The reviews like user generated data surround the emotional burst which after reading a book is generated or sometimes after watching a movie. These feelings earned from audits can be utilized in suggesting things from amusement space with comparable feelings. Till now much work has been finished utilizing feelings as discrete highlights in recommender framework.

It is hard to know how clients interpret their judgment of different item includes into evaluations, and much harder to comprehend why a thing may not be prescribed by the framework. Notwithstanding when diverse clients dole out a comparable generally speaking item evaluating, each may think about various highlights of the item as is apparent from audit remarks. Mining unequivocal client feelings on item includes from audits can help surmise client inclinations at a better granularity.

1.5 Proposed Work

By the analysis which is conducted carefully regarding the challenges and concerns, a fresh perspective is suggested by us for the recommendation generation across various domains by the emotions tapping which are generally seen to be encapsulated inside the content which is user generated and textual like feedback and reviews, articles and blogs along with comments and better performance is expected in the issues mitigating such as scalability, cold start and thinness of data. Using emotions as discrete features in the system of recommender till now much work has been done. As auxiliary domains in this work we generally delve deeper to make use of user sentiments for the recommender system cross domain enhancement. In identification of products with like characteristics for recommendation the use of some emotional aspects such as joy, love, surprise and sadness are often effective.

1.6 Our contributions are summarized as follows:

- Innovative scheme is developed by us who is to make the profile of sentiments regarding the product and users inside the source and domains targeted by feedback and review system along with ratings. To shift the informational sentiments of the reviews of users and they can be correlated them across various domains to offer better recommendations for the products in the target domains to the new user.

- For the collective factorization and doing it to the matrices rating for the targeted as well as auxiliary data, initial latent factors are obtained by it for the utilizations and products in a sequence.
- About a source domain to make use of user sentiment information, for example, if one is interested to watch a movie which is thrilling or horror and generalization of this information to the targeted domain can be made for the suggestion of books which are related to thrill and horror.
- The sentiment profile enrichment and this is done by multiple media tapping such as captions, pictures, videos and other factors which to a product are associated.
- Cold start problem solution with the information transfer based on neighborhood is supposed to be utilized with the system to convey the new user so as to suggest product on the ratings scale which similar users have provided and carrying the information which is similar.
- The sparseness of targeted data is to be solved, factorization of the efficient matrix methods are utilized in such a way that the common knowledge which is more useful is transferred from the auxiliary to the targeted domain.
- Initial latent factors construction making use of a likeness, graphs of these latent factors and alikeness are regarded as bridge implicit for the shift or implication of information which is more useful from the domain source to the target.
- Using movies and books as representative domains to mention our approach, making use of movies set of data in imbd.com to depict the domain of the movie and the book which is crossing the dataset in amazon.com for the book domain representation. In none of the cases, the way is applicable essentially to any other pair for the source domain and targeted domain in the case where choices made by humans are emotionally driven such as video games, television show or any movie, a music etc.
- To get an approach which is based on systematic testing for the user verification that the degree of acceptance or rejection is reflected by the ratings especially for the content this is emotional for the products. On the basis of such assumptions, the suggested system performance is evaluated by us. To validate the performance of the suggested method experimental outcomes are performed for the target domain

prediction making use of many metrics like RMSE, MAE etc. with similarities of Pearson or cosine.

1.7 Related Works

The systems of algorithmic recommenders try to forecast the products which will be liked by the target users on the basis of data or relevant information regarding the preference of users and the larger communities' preferences. Since it has been in use for more than a decade, the system users as well as investigative workers still have some argument about the performance of such impersonal recommender systems and the way they perform actually along with the recommenders which are humans. The race of who better predicts [16]? From an online study the outcomes contrast online system of recommenders and humans recommenders. The performance of algorithmic predictions of the MovieLens is attempted by the authors with the user profile based recommendations by the users of active MovieLens. It has been found by the authors that collaborative algorithm filtered the human outperformance on average despite of the fact that people outperformed the substantiality of the system as well as humans and net outperformed the system on tasks related to certain predictions.

One of the main challenges in the track for weekly recommendation is the model designing which is able to cope with the user characteristics or product characteristic dependent on time. in order to resolve the thinness of data problem in the process of collaborative filtering, it has been introduced by the research workers that viable approach is the transfer learning to make better use of auxiliary data. In collaborative filtering most previous learning works and focus has been kept on the points of exploiting according to ratings such as number based ratings, binary likes or dislikes ratings, stars etc. On the other hand, in the systems of real world recommenders, there is lesser chance that users will be able to rate items with precision and they often are unwilling or unlikely to do so. In comparison, non preference data can be turned by the practitioners for the estimation of a range and distribution of rating for the preference of user for an item. The name given to range or rating distribution is uncertain rating as it mentions the spectrum of rating with uncertainty despite of score based on point wise accuracy. The transfer learning process [17] for the collaborative filtering with the help of model of rating matrix generation, the solutions have been mentioned by the authors for the track of their weekly recommendations for CAMRA 2010 the social network

tracking challenge. In direction of this aim, two general approaches have been considered and compared by the authors, one approach is data weighting, and the other approach is a time aware modeling. By the extension of well known model of neighborhood or matrix factorization both these approaches can be implemented. For the track of social network, two extensions of the matrix factorization models have been developed by the authors and compared as well so that to the social network body they can be added and they are named as CMF which is an abbreviation of collective matrix factorization and the NRMF which stands for the regularized matrix factorization.

For the enhancement of models of matrix factorization, in the neighborhood adaptation and models of matrix factorization for the recommendations which are aware of context [18]. To weekly recommendation track and CAMRA 2010 social network track the authors describe their solutions. The CMF which stands for the collective matrix factorization and the NRMF which stands for the network regularized matrix factorization when these are contrasted. It has been shown in the outcomes of experiments that temporal information when used can result in the much more improvement. The same use on the MRMF can result in slight improvement by the combination of social networking inside the data rating.

In the music information retrieval MIR as the social tagging is becoming much more popular, the users are allowed to tag their favorite music items such as singer, album or any song. To MIR these social tags are highly valuable since multiple faced source of information is seen in them regarding the mood of users, style and opinion and the tool. Authors examine the problems in this paper for the recommendation on the personalized music based on the tags given at social media, it has been suggested by the authors that social tagging data and their modeling with tensors of three order which is able to get hold of three way correlations between music, users and tags. In the recommendation of the Musibox for the personalized music on the basis of social tags and cubic analysis [19], it has been suggested that in this model the latent structure has been developed with more HOSVD which stands for the higher order singular value decomposition. This also offers more accuracy and recommendations which are personalized, for example to the particular preference of user they are adapted. In the social tagging data to address the existing sparseness and to improve the recommendation quality further, it has been suggested by the research workers to make better models with the scheme of tag propagation which makes use of values of similarity evaluated between the items of music based on the features of their audio. As an outcome

of this the model which is suggested becomes highly effective when both information are combined regarding the features of audio as well as social tags. For the model suggested the performance is experimentally examined with the last fm real data. It is indicated in the outcomes that the suggested model is superior and have much better suggestive approach as compared to other methods which exist which overcome the cubic linkage in social tagging data that are inherent. Furthermore, it has been suggested in the results that social tagging data when combined with some other features such as audio is given much more preference as compared to the independent utilization of the former.

The recommendation of context awareness is sought to make better performance based recommendations by the information exploiting along with the recommender systems which make use of conventional user item matrix. To satisfy a specific purpose the recommendation should strive usually. In the mood of mining the movie which is similarly specific with the factorization matrix for the recommendation based on the awareness of a context. [20]. A novel movie similarity measure has been presented by the authors which to the movie property is specific by the challenging demand. i.e., mood for a movie. By a joint matrix factorization their measure is further exploited for the recommendation. It is experimentally validated by the authors that the suggested algorithmic effectiveness in the exploitation of mood specific movie is based on the similarity for the suggestion according the metrics of evaluation at several scales and showing that it is able to outperform some approaches which are state to art. Specifically speaking, the similarity of the mood specific movie is shown to be rather advantageous as compared to the similarity based on mood for the sake of recommendation which is mood specific.

In context awareness systems of recommenders since we know that mood has proven to be a significant contextual feature. In the ways which are based on mood for the collective filtering [21] it has been suggested by the authors that two new ways or approaches are there which are based on mood and are known as CF which stands for the hybrid collaborative filtering, and this is to ensure that accuracy of performance is improved and user is more satisfied by CARS making use of the emotional context. The traditional based CF is first described by the authors as the approach which is base lined and a new method is also suggested which is based on mood for the CF user for the detection of each emotion and user preference. Making use of multiple step the authors suggested the two hybrid approaches on

this basis and nearest neighbors are researched and forecasted fusion of rating is made respectively from the two strategies. Experimental comparisons are performed by the authors for the prediction of approaches mentioned above regarding the dataset of the Movie pilot for the movie recommendation on the basis of context awareness. (CAMRa20100. Both hybrid approaches offer better performance and much more improvement as suggested in the outcomes.

Likewise in "The part of client state of mind in film recommendations"[22] creators analyze how clients' mind-set can affect their examination of motion pictures in various classifications, which thus can help "educate recommender arrangement of grabbing motion pictures that are fitting for clients in various temperament. In particular, creators did two examinations. The principal comprises of a progression of client concentrates to look at client mind-set and motion picture evaluations to answer questions like: will a client in a more positive disposition tend to rate a lighthearted comedy higher? Will a client in a more anxious disposition tend to rate an activity motion picture higher? At that point, drawn upon the outcomes from the principal consider, creators change the conventional shared separating based suggestion approach by infusing client state of mind and proposed a disposition mindful synergistic sifting approach. Observational investigations exhibit that the disposition mind-ful suggestion approach performs superior to anything customary one that does not think about inclination.

Based on framework for the recommendation of the cross domain in a generic system [23]. It has been presented by the authors that investigative task on the development and designing of the knowledge based on generic model description is made on the semantic networks. The exploitation of knowledge and integrating it is aimed on the various domains to offer the recommendations of cross domain. To be more specific, an approach is presented by the authors which extract the data automatically regarding various domains such as music or either architecture which at the repositories of the linked data is available. This also allows the concept of linkage in two domains directed to acyclic graph by means of weighted directions and weight spreading is performed on this graph which is to recognize all the things or places of interest in the source domain.

The Setting mindful "music recommender frameworks are fit to propose music things mulling over logical conditions, for example, the client disposition or area, that may impact the client inclinations at a specific minute. "To address this issue in " Location-adjusted music proposal

utilizing labels" [24] "a specific sort of setting mindful suggestion undertaking — choosing music content that fits a position of intrigue (POI) is utilized creators have utilized passionate labels appended by a clients' populace to both music and POIs. In addition, creators have considered an arrangement of likeness measurements for labeled assets to build up a match between music tracks and POIs. Keeping in mind the end goal to test their speculation, i.e., that the clients will figure that a music tracksuits a POI when this track is chosen by their approach, creators have outlined a live client analyze where subjects are over and over given POIs and a choice of music tracks, some of them coordinating the exhibited POI and some not. The consequences of the trial demonstrate that there is a solid cover between the clients' determinations and the best coordinating music that is suggested by the framework for a POI. "

Again change in Transfer is proposed In “Transfer learning in communitarian separating with unverifiable evaluations" [25], "the creators propose a productive exchange learning answer for shared sifting, known as an exchange by integrative factorization (TIF), to use such helper dubious appraisals to enhance the execution of suggestion. Specifically, creators incorporate assistant information of questionable evaluations as extra requirements in the objective framework factorization issue and take in a normal rating an incentive for each dubious rating naturally. The upsides of their proposed approach incorporate the proficiency and the enhanced adequacy of shared sifting; demonstrating that joining the helper information of indeterminate evaluations can truly bring an advantage. Exploratory outcomes on two motion picture proposal assignments demonstrate that their TIF calculation performs altogether better finished a cutting edge non-exchange learning strategy. "

On the models which are based on the dimensional emotions and social based tags [26], an emotional computational model is presented by authors which are based generally on the social labels. On the automatically produced lexicon the model is generally build upon and it also explains the emotions by terms such as synonyms and antonyms. To the multiple domains which are emotion specific are linked and from the entertainment systems of social tagging the folksonomies are suggested. For the folksonomies making use of cross domain, the research workers suggest a lot of methods which are able to transform tag based product profiles automatically into the profiles based on the orientation of emotions. Authors also reported the outcomes in order to make their model validate from the study of user showing higher degree of precision for the methodology used so as to infer the evoked emotions by

the products in the domains such as movies or music and as a consequence from the evaluations taking place offline show much more accuracy in the recommenders which are based on models and their improvement which is added into the information for the extraction of items.

The tags are exploited by the recommendations based on the systems along with text based descriptions or available ratings for the product in a single domain so as to suggest multiple domains and the product to be preferred. By the handling of unnannotated information and unstructured information handling is always a big challenge. For deducing structure in such corpora of data the topic modeling offers a common method. In the recommendation of the cross domain based on semantic clustering" [27], it has been introduced by the research workers that common latent concept for the semantic space as well as multiple domain spanning, for the distinct domains making use of modeling topic for the semantic clustered vocabularies. To use the relationship which is determined explicitly semantic is the main intuition here as there is non identical linkage but there is a possibility of equivalent based on semantics, other domain words and vocabularies so as to get through the association which is obtained in the domains which are distint. The WordNet which is popular one is generally ontology dependent is utilized for the measurement of relatedness semantically between the words used in the textual form. It has been shown in the outcomes of experiments that in the precision of user preference prediction there is much more improvement for the singular domain especially in the case where other domains are given more preference.

In the system of recommender the new user problem is always a big challenge and no solution which is unique one is able to be applied in any situation or domain. The analysis of viable solutions is made by the research workers in this paper to the collaborative filtering for the new user problems which on the personality information of the user are exploited. (a) the CF which is based on personality and is able to enhance the model of prediction recommendation by adding the information and personality data, (b) active learning which is based on personality and makes use of information regarding personality for the preference identification in the targeted data in the domain recommendation which from the user is to be elicited, (c) based on the personality the recommendations across various domains which makes use of personalized information to make better preferences of user data from the auxiliary domains and for the compensation of the lack of user preference they can be used simply in the domains targeted. In the process of collaborative filtering the alleviation of new

user problem and this is done by the exploitation of the information regarding personality. [28]. The effectiveness of these models can be benchmarked by the authors on the datasets which are larger spanning many domains such as books, movies or either music and methods of personality awareness are achieved in the improvement of their performance ranging from 6 to 94% for all the new users of the system. In the mean time the novelty of the system from 3 to 40% is increased in respect to the baseline popularity which is none personalized. The limitations of the approach are also discussed by the authors along with situations in the methods suggested which can be applied in much better way. It also offers the guideline for the research workers and other practitioners working in the same field. For the ambient intelligence environment there has been an overwhelming of the users by the number at social media which is considered to be larger one and hence to identify the needs of users, the social media can be used and tailored. At the current time it is regarded as a genuine question to be asked.

For the environments which have ambient intelligence, their users have been seen to provide overwhelming response to the available number of social media and in this way the number of questions can be asked and needs of the users can be identified and tailored at social media. In the social media services the recommendation of tag based personalization is adapted by the research workers to measure the Katz proximity and this is to be used in the tagging system at social media so that from ambient environment the users can be helped and they can find the media which suits their own personal needs. The methods presented in the model shows the ternay relationship existing in between the users, as weighted the resource and tag with the graphs being undirected and tripartite, the Katz proximity measure is then applied to the graph. On the two datasets the experiments performed are implemented and they are contrasted with the algorithms which are state to art. It has been proven by the outcomes of experiments that Katz algorithmic adaptation with the tripartite framework has better improvement and also it ranks successfully and the results are obtained as per the interest of users.

To help people make choices and take decisions is an important function of recommender system. Reviews and comments that are written online by users after watching movies, reading books, listening to music or purchasing a specific item, are important sources of user generated data, which can be utilized for decision making using recommendations given to user by the recommender system. When the available data in one domain is sparse, data

(content and ratings) from other domains can be used to make cross domain recommendations. In **" On using reviews and comments for cross domain recommendations and decision making"** [30] authors have used this content information as well as ratings of both domains where there is no user-item overlap between the given domains for cross domain recommendations. User generated content (reviews and comments) crawled from web requires topic modeling to discover the latent thematic structure in the corpora of both domains. Since the topics in both domains are dissimilar, authors have compared various approaches based on semantic space of corpus and knowledge based methods for finding cross domain recommendations. Experimental results show improvement in precision in recommendations over existing approach based on semantic clustering.

In the recent time while reviewing the profiles based on emotions for the recommendation across various domains it has been suggested by the authors that for the recommendation generation a fresh perspective across various domains cab be tapped by the encapsulated emotions inside the textual content generated which include blogs, comments, feedbacks and reviews. Between various entertaining domains these emotions are thought to serve as forceful links based on social psychology and they have potential to solve the problems regarding the cold start. The CDR scheme used by them is rich in the lexicon emotions which can analyze the online emotional content which users express in the targeted domains and source and emotional based profile is generated for products in various domains. In the meanwhile the collaborative filtering is applied so as to match these profiles so that targeted domains are recommended by them. Some the basis of CDR emotion based model they are illustrated and book domains are used for the case study. On the datasets of Movielens and Bookcrossing the outcomes of experiments indicate that 28.9 is the yield which measures F1 for the marked improvement of 71.1 percent in contrast to the approach based on topic modeling for the entertainment domains of CDR.

1.8 Methodology

The approach or the system of cross domain recommendation is discussed in this paper making use of the knowledge regarding transfer from the sentimental analysis. The system of recommender which is suggested here contains various modules which are called processing

of data, analysis of sentiments, building models of CDR and evaluation and computational similarity.

1.8.1 Data Processing

At the ecommerce sites and various social media platforms it is seen that users express their feedback and opinion openly regarding the product. All the reviews are extracted by this module regarding a particular movie and concentrates on the feedback which is based on the reviews in the form of textual content regarding a movie. The informal style in used by the users when they write a review. Corpus is formed due to the preprocessing of such reviews and irrelevant content is removed and the relevant content is left behind. The following steps are involved in the preprocessing of the review corpus:

1. The process known as tokenization whereby the word constituents from the review corpus are taken and collected as tokens.
2. Word removal stoppage whereby the common and familiar words from the token such are a, an, is, at, which are regarded as irrelevant for the analysis of emotions from the corpus are removed.
3. Lowercase conversion whereby the review corpus as a whole is altered to the lowercase and this is done to maintain the token coherence.
4. The process of stemming whereby the root words are taken from the token maps which as shock, shockingly, shocked etc.
5. Compression of word especially in review writing it is seen that users tend to prolong it such as by writing coooool instead of cool, happppppppy instead of happy etc. such elongated words are identified and compressed in this step.

1.8.2 Sentiment Analysis

Opinion Mining [32] is the way toward identifying the relevant extremity of twitter posts [33], News, Blog and some other content articles. At the end of the day, it decides if a bit of composing is sure, negative or impartial. An elective term is assessment mining, as it infers the conclusion, or the disposition of tweet. A typical utilizes case for this innovation is to find how individuals feel about a specific subject. Opinion examination is one field of NLP [33] which is pulling in incredible consideration from scientists. News and online journals are generally great wellsprings of information for supposition examination, wherein individuals

can express their musings and feelings on such gatherings. The handling framework incorporates an assessment recognizable proof stage first and foremost. It distinguishes the different elements in the twitter posts and connects communicated assessments with each significant element. The second stage includes feeling collection and scoring, where every element is scored in respect to others in a similar class. This will evaluate the assessments for every substance. The principle focal point of the framework is distinguishing semantic introduction of words in the primary stage, utilizing WordNet [34]". A mix of POS labeling, WordNet and SentiWordNet [35] and an expansive exhaustive vocabulary particularly for target sentences will consider opinion examination to be performed for target sentences.

1.8.3 Model Building

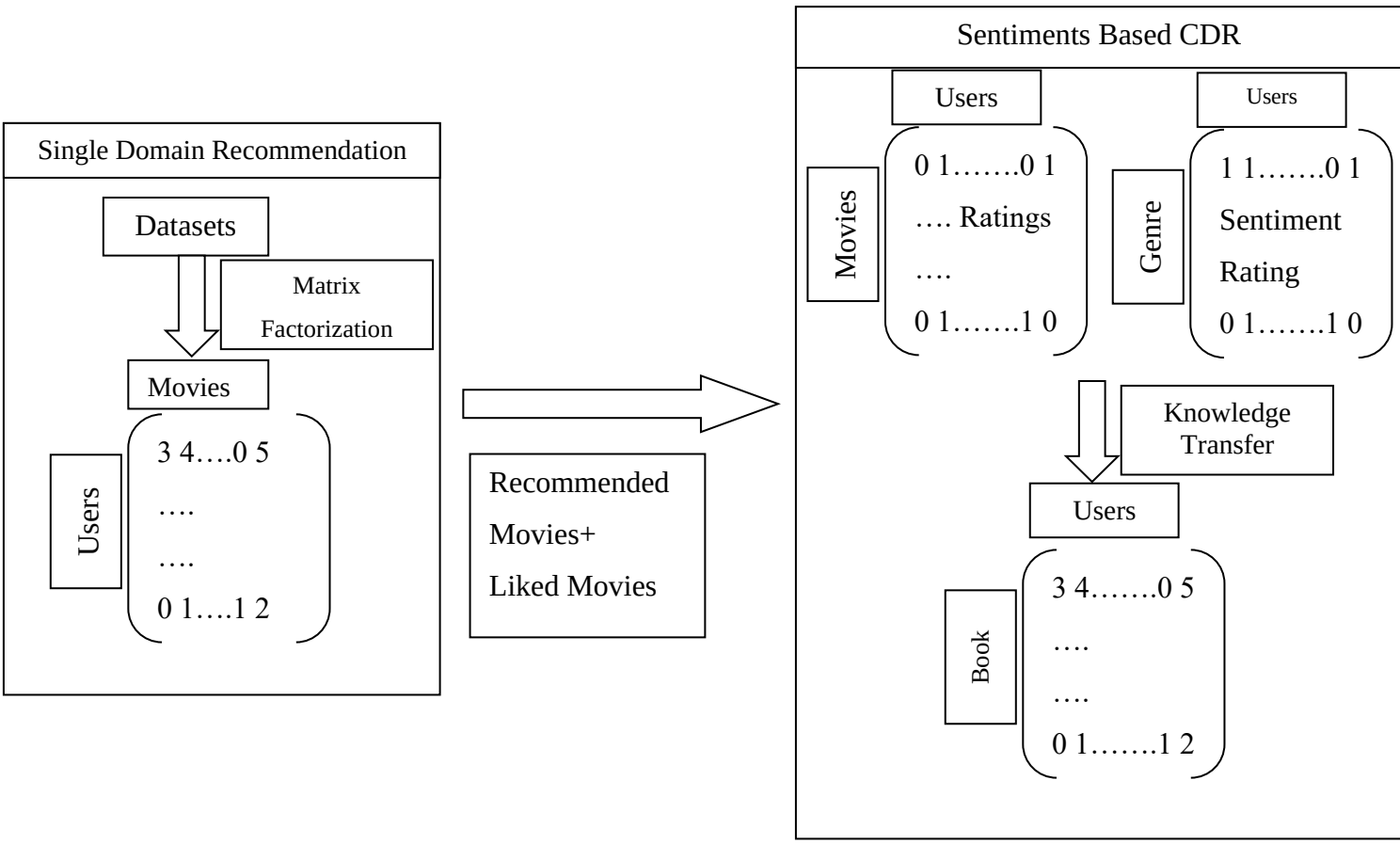


Figure 4 : Proposed Model

1.8.4 Similarity Computation.

Between the items or users the computation of similarity is always an important step in the collaborative filtering based on memory for all the algorithms. For the algorithmic CDR the similarity computation idea between the i item and j item is to be investigated at first especially for the users who have seen to rate both the items and similarity computation is then applied in order to find out the ratio of similarity $w_{i,j}$, for the users between two rated items. For the algorithmic CDR figure, the similarity is computed at first $w_{u,v}$, between the u and v users who have been seen to rate the two similar items. To compute the similarity or weight between users or products there are a lot of ways.

1.8.4.1 Correlation-Based Similarity

The similarity in this case is $w_{u,v}$ in between u and v users or $w_{i,j}$ for the i and j item, the Pearson correlation is used to make measurements or similarities based on correlations.

Pearson correlation [36] the extent is measured to which linear variables are related to each other, the correlation between Pearson in between u and v users is mentioned below:

$$w_{u,v} = \frac{\sum_{i \in I}(r_{u,i} - \bar{r}_u)(r_{v,i} - \bar{r}_v)}{\sqrt{\sum_{i \in I}(r_{u,i} - \bar{r}_u)^2}\sqrt{\sum_{i \in I}(r_{v,i} - \bar{r}_v)^2}}$$

In this case, the i∈ I summations are over the items that both the users u and v have rated and $\bar{r}_u$u is the average rating of the co-rated items of the uth user. In an example in Table 4, we have $w_{1,5}$ = 0.756.

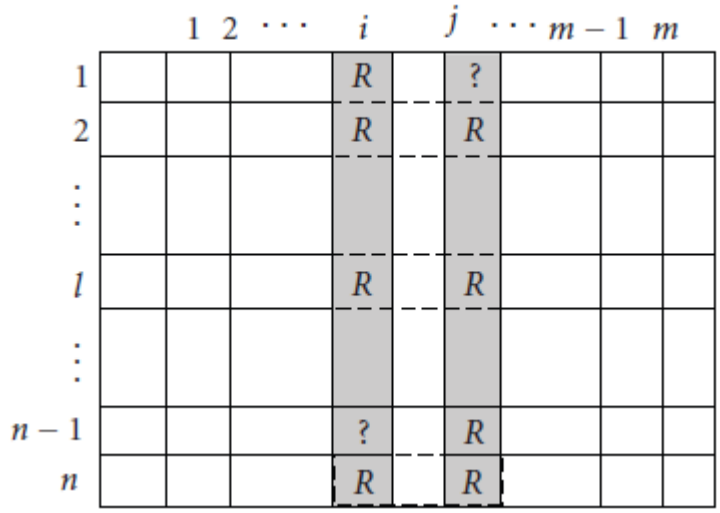


Figure5: Item-based similarity ($w_{i,j}$) calculation based on the co rated items i and j from users 2, l and n.

For the item-based algorithm, denote the set of users u ∈U who rated both items i and j, then the Pearson Correlation will be

$$w_{u,v} = \frac{\sum_{u \in U}(r_{u,i} - \bar{r}_i)(r_{u,j} - \bar{r}_j)}{\sqrt{\sum_{i \in I}(r_{u,i} - \bar{r}_i)^2}\sqrt{\sum_{i \in I}(r_{v,j} - \bar{r}_j)^2}}$$

where $r_{u,i}$ is the rating of user u on item i, $\bar{r}_i$ is the net for both item ratings given by the users as mentioned in the diagram.

On the item basis some variations and correlations based on users by Pearson are supposed to be estimated.

1.8.4.2 Vector Cosine-Based Similarity.

There is some sort of a likeness in between two documents which one can estimate by the document treatment of vector for the frequencies and cosine computation for the formed angle for the frequency vectors. In collaborative filtering this formalism can be adopted which makes use of products rather than using documents and use ratings despite of frequency words. [37].

Speaking formally, in any case where the m × n is the R item matrix, between the two items the similarity i and j as cosine of n dimensional vectors are defined which correspond to ith and jth matrix R columns.

Between the items i and j the similarity of vector cosine is represented as:

$$w_{i,j} = Cos(\vec{i}, \vec{j}) = \frac{\vec{i}, \bullet \vec{j}}{\|\vec{i}\| * \|\vec{j}\|}$$

In this case "•" represents two vector's dot product and in order to get the similarity computation as desired for the items n and similarity matrix n x n can be computed as well. in this case if $\vec{A}$= {x_1, y_1}is a vector A, vector $\vec{B}$ = {x_2, y_2}, between A and B the similarity in the vector cosine between A and B is represented as:

$$w_{A,B} = Cos(\vec{A}, \vec{B}) = \frac{\vec{A}, \bullet \vec{B}}{\|\vec{A}\| * \|\vec{B}\|} = \frac{x_1 x_2 + y_1 y_2}{\sqrt{x_1^2 + y_1^2}\sqrt{x_2^2 + y_2^2}}$$

In the real scenarios, various customers are supposed to make use of scales of different ratings which cannot be taken into account by the vector cosine similarity. This drawback in order to be addressed is adjusted according to the similarity cosine which by the corresponding user is subtracted and from each of the co rated pair the average is taken. The similarity of the adjusted cosine contains the formula which is similar to the Pearson correlation. The cosine similarity of the Pearson correlations as a matter of fact performs the normalization of same sort as the rating by users according to the behavior of rating. We may get the negative values with the correlation of Pearson but this is not the case with similarity cosine keeping in mind that we have a rating scale of n point.

1.9 Datasets Used

From Movie Lens the dataset has been obtained for the experimental purpose [38] according to the website (http://movielens.org) by the Group lens research this is collected and made available [39]. 100,000 ratings are there in the MovieLens dataset which range from one to five with the users equal to 943 while films equal to 1682. The research presented here suggests that the users which are 1100 are there with ratings above 30. On the other hand the dataset regarding the library thing comprises of the rating data set of the books making up ratings of over 1.1 million on the scale 1-10 by 7279 users on 37,232 books. From the top to recommend books we have selected 5000 tags. In the form of two dimensional matrixes the user item dataset is made available.

1.10 SYSTEMREQUIREMENTS

PROCESSOR	: Intel CORE i5
SYSTEM BUS	: 64 Bit
RAM	: 6 GB
HDD	: 500 GB
MODEM	: 128 KBPS

1.10.1 Evaluation Metrics

By comparing the recommendations to the test set of user ratings which are known the quality of the system of recommender can be evaluated. By using predictive accuracy metrics these systems are typically measured and in this case the ratings are predicted and compared directly to the user ratings which are accurate. In the literature the metric which is used most commonly is known as the Mean Absolute error (MAE)[41][42] – in between the predicted ratings and actual ratings it is defined as the average absolute difference and is represented as:

$$MAE = \frac{\sum_{\{u,i\}} \left| p_{u,i} - r_{u,i} \right|}{N}$$

In this case $p_{u,i}$ is the rating which is predicted for the u user and I item, the actual rating is $r_{u,i}$, while the total rating number is N in the set of test.

A metric which is related and commonly used is the RMSE which stands for root mean squared error, it lays more emphasis on the errors which are absolutely larger ones and are represented as:

$$RMSE = \sqrt{\frac{\sum_{\{u,i\}}\left(p_{u,i} - r_{u,i}\right)^2}{N}}$$

All the items are equally treated by the predictive accuracy metrics. On the other hand, for the systems of recommenders used most commonly, the main concern is to deal with predicting accuracy for the products which customers will prefer. For example, the research workers sometimes see the suggesting as a good tool for prediction, i.e. the products having better and higher ratings Vs poorly or badly rated products.

1.11 References

[1] Viktoratos, I., Tsadiras, A., & Bassiliades, N. (2018). Combining community-based knowledge with association rule mining to alleviate the cold start problem in context-aware recommender systems. Expert Systems with Applications, 101, 78-90.

[2] Melville, P., & Sindhwani, V. (2011). Recommender systems. In Encyclopedia of machine learning (pp. 829-838). Springer, Boston, MA.

[3] Al Mamunur Rashid, S. K. L., Karypis, G., & Riedl, J. (2006). ClustKNN: a highly scalable hybrid model-& memory-based CF algorithm. Proceeding of WebKDD, 2006.

[4] Shani, G., & Gunawardana, A. (2011). Evaluating recommendation systems. In Recommender systems handbook (pp. 257-297). Springer, Boston, MA.

[5] Sarwar, B., Karypis, G., Konstan, J., & Riedl, J. (2002, December). Incremental singular value decomposition algorithms for highly scalable recommender systems. In Fifth International Conference on Computer and Information Science (pp. 27-28). Citeseer.

[6] Koren, Y., Bell, R., & Volinsky, C. (2009). Matrix factorization techniques for recommender systems. Computer, (8), 30-37.

[7] Peterson, L. E. (2009). K-nearest neighbor. Scholarpedia, 4(2), 1883.

[8] Wilson, D. C., Smyth, B., & Sullivan, D. O. (2003). Sparsity reduction in collaborative recommendation: A case-based approach. International journal of pattern recognition and artificial intelligence, 17(05), 863-884.

[9] Wang, X., Peng, Z., Wang, S., Philip, S. Y., Fu, W., & Hong, X. (2018, May). Cross-domain recommendation for cold-start users via neighborhood based feature mapping. In International Conference on Database Systems for Advanced Applications (pp. 158-165). Springer, Cham.

[10] Cheng, W., Yin, G., Dong, Y., Dong, H., & Zhang, W. (2016). Collaborative filtering recommendation on users' interest sequences. PloS one, 11(5), e0155739.

[11] Pan, W., Xiang, E. W., Liu, N. N., & Yang, Q. (2010, July). Transfer Learning in Collaborative Filtering for Sparsity Reduction. In AAAI (Vol. 10, pp. 230-235).

[12] Pan, W., & Yang, Q. (2013). Transfer learning in heterogeneous collaborative filtering domains. Artificial intelligence, 197, 39-55.

[13] Pan, S. J., & Yang, Q. (2010). A survey on transfer learning. IEEE Transactions on knowledge and data engineering, 22(10), 1345-1359.

[14] Pan, W., Chen, L., & Ming, Z. (2018). Personalized recommendation with implicit feedback via learning pairwise preferences over item-sets. Knowledge and Information Systems, 1-24.

[15] Pan, W., Yang, Q., Duan, Y., Tan, B., & Ming, Z. (2017). Transfer Learning for Behavior Ranking. ACM Transactions on Intelligent Systems and Technology (TIST), 8(5), 65.

[16] Tang, T. Y., & Winoto, P. (2008, July). Mood and recommendations: On non-cognitive mood inducers for high quality recommendation. In Asia-Pacific Conference on Computer Human Interaction (pp. 94-103). Springer, Berlin, Heidelberg.

[17] Li B, Yang Q, Xue X (2009) Transfer learning for collaborative filtering via a rating-matrix generative Model. Proc of the 26th Annual International Conference on Machine Learning: 617–624

[18] Liu, N. N., Cao, B., Zhao, M., & Yang, Q. (2010, September). Adapting neighborhood and matrix factorization models for context aware recommendation. In Proceedings of the Workshop on Context-Aware Movie Recommendation (pp. 7-13). ACM.

[19] Nanopoulos, A., Rafailidis, D., Symeonidis, P., & Manolopoulos, Y. (2010). Musicbox: Personalized music recommendation based on cubic analysis of social tags. IEEE Transactions on Audio, Speech, and Language Processing, 18(2), 407-412.

[20] Shi Y, Larson M, Hanjalic A (2010) Mining mood-specific movie similarity with matrix factorization for context-aware recommendation. Proc of the workshop on context-aware movie recommendation, p 34–40

[21] Wang, L., Meng, X., Zhang, Y., & Shi, Y. (2010, September). New approaches to mood-based hybrid collaborative filtering. In Proceedings of the workshop on context-aware movie recommendation (pp. 28-33). ACM.

[22] Winoto P, Tang TY (2010) The role of user mood in movie recommendations. Expert Syst Appl 37(8): 6086–6092

[23] Fernández-Tobías I, Cantador I, Kaminskas M, Ricci F (2011) A generic semantic-based framework for cross-domain recommendation. Proc of the ACM 2nd International Workshop on Information Heterogeneity and Fusion in Recommender Systems: 25–32

[24] Kaminskas M, Ricci F (2011) Location-adapted music recommendation using tags. Proc Springer Berlin Heidelberg. International Conference on User Modeling, Adaptation, and Personalization, p 183–194

[25] Pan W, Xiang EW, Yang Q (2012) Transfer learning in collaborative filtering with uncertain ratings. In AAAI

[26] Fernández-Tobías I, Cantador I, Plaza L (2013) An emotion dimensional model based on social tags: Crossing folksonomies and enhancing recommendations. Proc Springer

Berlin Heidelberg. International Conference on Electronic Commerce and Web Technologies, p 88–100

[27] Kumar A, Kumar N, Hussain M, et al (2014) Semantic clustering-based cross-Domain recommendation. Proc Computational Intelligence and Data Mining (CIDM):137–141

[28] Fernández-Tobías I, Braunhofer M, Elahi M et al (2015) Alleviating the new user problem in collaborative filtering by exploiting personality information. User Model User-Adap Inter 1–35

[29] Rawashdeh M, Alhamid MF, Alja'am JM et al (2016) Tag-based personalized recommendation in social media services. Multimedia Tools and Applications 75:13299. doi:10.1007/s11042-015-2813-0

[30] SaraswatM, Chakraverty S,Mahajan N, Tokas N (2016) On using reviews and comments for cross domain recommendations and decision making. In: Computing for Sustainable Global Development (INDIACom), 2016 3rd International Conference on IEEE, p 3656–3659

[31] Chakraverty, S., & Saraswat, M. (2017). Review based emotion profiles for cross domain recommendation. Multimedia Tools and Applications, 76(24), 25827-25850.

[32] Pang, B., & Lee, L. (2008). Opinion mining and sentiment analysis. Foundations and Trends® in Information Retrieval, 2(1–2), 1-135.

[33] Speriosu, M., Sudan, N., Upadhyay, S., & Baldridge, J. (2011, July). Twitter polarity classification with label propagation over lexical links and the follower graph. In Proceedings of the First workshop on Unsupervised Learning in NLP (pp. 53-63). Association for Computational Linguistics.

[34] Miller, G. A. (1995). WordNet: a lexical database for English. Communications of the ACM, 38(11), 39-41.

[35] Baccianella, S., Esuli, A., & Sebastiani, F. (2010, May). Sentiwordnet 3.0: an enhanced lexical resource for sentiment analysis and opinion mining. In Lrec (Vol. 10, No. 2010, pp. 2200-2204).

[36] Benesty, J., Chen, J., Huang, Y., & Cohen, I. (2009). Pearson correlation coefficient. In Noise reduction in speech processing (pp. 1-4). Springer, Berlin, Heidelberg.

[37] Tata, S., & Patel, J. M. (2007). Estimating the selectivity of tf-idf based cosine similarity predicates. ACM Sigmod Record, 36(2), 7-12.

[38] Miller, B. N., Albert, I., Lam, S. K., Konstan, J. A., & Riedl, J. (2003, January). MovieLens unplugged: experiences with an occasionally connected recommender system. In Proceedings of the 8th international conference on Intelligent user interfaces (pp. 263-266). ACM.

[39] Resnick, P., Iacovou, N., Suchak, M., Bergstrom, P., & Riedl, J. (1994, October). GroupLens: an open architecture for collaborative filtering of netnews. In Proceedings of the 1994 ACM conference on Computer supported cooperative work (pp. 175-186). ACM.

[40] Lu, C., Park, J. R., & Hu, X. (2010). User tags versus expert-assigned subject terms: A comparison of LibraryThing tags and Library of Congress Subject Headings. Journal of information science, 36(6), 763-779.

[41] Willmott, C. J., & Matsuura, K. (2005). Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance. Climate research, 30(1), 79-82.

[42] Chai, T., & Draxler, R. R. (2014). Root mean square error (RMSE) or mean absolute error (MAE)?–Arguments against avoiding RMSE in the literature. Geoscientific model development, 7(3), 1247-1250.

[43] Daniel Kluver, "What is a recommender system", 2015 [Online Available] https://www-users.cs.umn.edu/~kluve018/what-is-a-recommender.htm

MANAV RACHNA
vidyanatariksha

Swami Vivekananda Library and Resource Centre
Manav Rachna University

SIMILARITY VERIFICATION CERTIFICATE

Title: Sentiment Enriched Cross Domain Recommender System

Researcher: Nikita Taneja Registration Number: 2K17-PFE-007

Supervisor: Dr. Hardeo Kumar Thakur Department: Computer Science

This is to certify that the above mentioned Short Synopsis/Thesis was scanned for similarity index through Turnitin Software on 27/10/21 with submission ID 1685431371 and the similarity index was found to be 10 %. The complete report running into 36 Pages, 9516 Word count with File Size 4001 KBis being forwarded to the Research Supervisor/DRC/DRB for review

Date 28/10/21

Swami Vivekananda Library and Resource Centre ★ Manav Rachna University ★ Faridabad
28/10/21

The complete report of the above Short Synopsis/Thesis has been reviewed and the similarity index is below the accepted norms as per the University Grants Commission (Promotion of Academic Integrity and Prevention of Plagiarism in Higher Educational Institutions) Regulations 2018. The Short Synopsis/Thesis may be considered for further processing

Research Scholar 28/10/21

Research Supervisor 28/10/2021

Nikita_Taneja_Short_Synopsis_27.10.2021

by Nikita 1 Taneja

Submission date: 27-Oct-2021 01:09PM (UTC+0530)
Submission ID: 1685431371
File name: Sentiment_Enriched_Nikita_Taneja_Short_Synopsis_27.10.2021.docx (129.4K)
Word count: 9516
Character count: 50209

Nikita_Taneja_Short_Synopsis_27.10.2021

ORIGINALITY REPORT

Excluding matches less than 14 words

After excluding ..... Research papers co-authored with the Research Supervisor, as annexed as .....

10%	3%	9%	1%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

Swami Vivekananda Library & Resource Centre, Faridabad * Manav Rachna University * Rayeshi Taneja 28/10/21

Harsh 28/10/2021

PRIMARY SOURCES

1	"Micro-Electronics and Telecommunication Engineering", Springer Science and Business Media LLC, 2020 Publication	4%
2	Xinghua Wang, Zhaohui Peng, Senzhang Wang, Philip S. Yu, Wenjing Fu, Xiaokang Xu, Xiaoguang Hong. "CDLFM: cross-domain recommendation for cold-start users via latent feature mapping", Knowledge and Information Systems, 2019 Publication	1%
3	link.springer.com Internet Source	1%
4	Xiaoyuan Su. "A Survey of Collaborative Filtering Techniques", Advances in Artificial Intelligence, 2009 Publication	1%
5	downloads.hindawi.com Internet Source	1%
6	documents.mx Internet Source	<1%

7	Submitted to University of Technology, Sydney Student Paper	<1%
8	Hui Qiao, Yangyang Liu, Xuewen Dong, Di Lu. "Personalized Recommendation for Cold-Start Users via Cluster-Level Latent Feature Model", 2019 International Conference on Networking and Network Applications (NaNA), 2019 Publication	<1%
9	www.citeulike.org Internet Source	<1%
10	"Cross-Domain Collaborative Recommendation by Transfer Learning of Heterogeneous Feedbacks", Lecture Notes in Computer Science, 2015. Publication	<1%
11	Submitted to Peking University Student Paper	<1%

Exclude quotes On

Exclude bibliography On

Exclude matches < 14 words